

Mirosława Sztemberg-Lewandowska

Akademia Ekonomiczna we Wrocławiu

ANALIZA CZYNNIKOWA ZMIENNYCH PORZĄDKOWYCH**1. Wstęp**

Klasyczna analiza czynnikowa bazuje na założeniu, że zmienne obserwowalne $(x_1, x_2, x_3, \dots, x_p)$ są metryczne. Wynika to z jej modelu:

$$x_i = a_{i0} + a_{i1}y_1 + a_{i2}y_2 + \dots + a_{ip}y_p + e_i \quad (i = 1, \dots, p), \quad (1)$$

gdzie: p – numer obserwacji,

x_i – i -ta obserwowalna zmienna,

$\mathbf{y} = (y_1, y_2, y_3, \dots, y_p)$ – wektor latentnych zmiennych,

e_i – reszta.

Dodatkowo zakłada się, że reszta ma rozkład normalny ze średnią 0 i wariancją σ_i^2 , ukryte zmienne są niezależne o rozkładzie normalnym $y_j \sim N(0,1)$ dla wszystkich j . Zatem x_i także jest zmienną metryczną.

W przypadku zmiennych porządkowych zakłada się, że obserwowalne niemetryczne zmienne są realizacjami fikcyjnych metrycznych ukrytych zmiennych (por. [Bartholomew i in. 2002; Konarski 2004]). Te ukryte zmienne są nieobserwowalne, jednak nie należy ich mylić z latentnymi zmiennymi. Obserwowalne porządkowe zmienne oznaczmy za pomocą $x_1^*, x_2^*, x_3^*, \dots, x_p^*$ oraz niekompletnie obserwowalne zmienne przez $x_1, x_2, x_3, \dots, x_p$. Każda zmienna x_i^* ma m_i kategorii oraz $m_i - 1$ progów, które oznaczmy za pomocą $t_{i1}, t_{i2}, t_{i3}, \dots, t_{i(m_i-1)}$. Progi te dzielą skalę ukrytej zmiennej x_i na m_i kategorii. Związek między niemierzalną zmienną x_i^* oraz mierzalną x_i jest następujący:

$$x_i^* = s \Leftrightarrow t_{i(s-1)} < x_i \leq t_{is} \quad (s = 1, 2, \dots, m_i), \quad (2)$$

gdzie: $t_{i0} = -\infty, t_{i1} < t_{i2} < \dots < t_{im_i-1}, t_{im_i} = +\infty$.

Dodatkowo zakłada się, że zmienna x_i ma średnią zero i wariancję równą jeden.

Model czynnikowy dla zmiennej x_i wygląda jak we wzorze (1), z tą różnicą, że w klasycznej analizie czynnikowej zmienna x_i jest bezpośrednio obserwowalna, a tutaj tylko częściowo poprzez x_i^* .

Analiza czynnikowa bazuje na macierzy korelacji między zmiennymi (por. [Bartholomew 1980]), w podejściu klasycznym dla zmiennych metrycznych wykorzystuje się korelacje liniowe Pearsona:

$$\rho_{x_1 x_2} = \frac{\sigma_{x_1 x_2}}{\sigma_{x_1} \sigma_{x_2}}, \quad (3)$$

gdzie $\sigma_{x_1 x_2}$ – kowariancja x_1 i x_2 , σ_{x_i} – wariancja x_i .

W odniesieniu do zmiennych porządkowych stosuje się korelacje polichoryczne. Rozkład dwóch zmiennych porządkowych kategoryalnych x_1^*, x_2^* przedstawiono w tab. 1.

Tabela 1. Tablica kontyngencji $m_1 \times m_2$

Wyszczególnienie	Pierwsza kategoria zmiennej x_2^*	Druga kategoria zmiennej x_2^*	...	Kategoria m_2 zmiennej x_2^*
Pierwsza kategoria zmiennej x_1^*	n_{11}	n_{12}	...	n_{1m_2}
Druga kategoria zmiennej x_1^*	n_{21}	n_{22}	...	n_{2m_2}
...	...	...	...	...
Kategoria m_1 zmiennej x_1^*	$n_{m_1 1}$	$n_{m_1 2}$	...	$n_{m_1 m_2}$

Źródło: opracowanie własne.

Niech $\phi_2(\cdot, \cdot, \rho)$ oznacza gęstość dwuwymiarowego rozkładu normalnego, π_{ij} – prawdopodobieństwo obserwacji $x_1^* = i \wedge x_2^* = j$ oraz (por. [Kukuk 1998; Olivares 2003]):

$$\pi_{ij} = P(x_1^* = i, x_2^* = j) = \int_{t_{1,i-1}}^{t_{1,i}} \int_{t_{2,j-1}}^{t_{2,j}} \phi_2(x_1^*, x_2^*, \rho) dx_1^* dx_2^*. \quad (4)$$

Korelacja polichoryczna ρ zmiennych x_1^*, x_2^* jest wyznaczana poprzez maksymalizację funkcji:

$$L(\rho, t_1, t_2) = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} n_{ij} \log \pi_{ij}, \quad (5)$$

gdzie n_{ij} jest wartością w i, j komórce tablicy kontyngencji (tab. 1).

Szczególnym przypadkiem zmiennych porządkowych są zmienne mierzone na skalach, które mają dwie kategorie (por. [Duffy i in. 1998; Ming Tang i in. 1999]), dla tych zmiennych wyznacza się korelacje tetrachoryczne, które estymuje się za pomocą równania (por. [Kubinger 2003]):

$$\int_{-\infty}^{t_1} \int_{-\infty}^{t_2} \phi(x_1^*, x_2^*, r) dx_1^* dx_2^* = \frac{n_{11}}{N}, \quad (6)$$

gdzie n_{11} – wartość w tablicy kontyngencji, $N = \sum_{i=1}^2 \sum_{j=1}^2 n_{ij}$, t_i – próg dla zmiennej

x_i^* , $\phi_2(x_1^*, x_2^*, r)$ – gęstość dwuwymiarowego rozkładu normalnego, r – korelacja tetrachoryczna.

W literaturze spotyka się zastosowanie tradycyjnej analizy czynnikowej dla zmiennych porządkowych. Celem artykułu jest zbadanie, czy takie podejście jest poprawne, jeśli tak, to czy dla wszystkich zmiennych porządkowych.

W artykule porównane zostaną wyniki analizy czynnikowej przeprowadzonej na podstawie korelacji liniowych Pearsona i korelacji polichorycznych dla zmiennych mierzonych na skali trzystopniowej i osobno dla zmiennych mierzonych na skali siedmiostopniowej. Na tej podstawie zbada się, czy podstawne jest zastosowanie korelacji liniowych Pearsona dla zmiennych porządkowych mierzonych na tych skalach.

Przedstawiony w artykule przykład empiryczny dotyczący rynku kawy ma na celu przedstawienie podobieństw i różnic w wynikach analizy czynnikowej po zastosowaniu obu korelacji.

2. Przykład empiryczny

Wykorzystane zostaną dane z badania ankietowego dotyczącego rynku kawy. Kawy oceniano ze względu na 9 cech: cena – B , aromat – D , wygoda w przygotowaniu – F , czystość przygotowania – G , estetyka opakowania – H , dostępność w sprzedaży – I , rzetelność informacji na opakowaniach – J , funkcjonalność opakowania – K , promocje – L . Zmienne zaprezentowano na skali trzystopniowej. Krawce skali określono w postaci antonimów. Ankietowany oceniał tylko znane mu kawy. W tabeli 2 podano kawy wraz z liczbą ankiet, w których były ocenione. Tabela 3 zawiera mediany ocen wszystkich kaw.

Zmienne zostały przedstawione na skali Likerta, zatem są to zmienne porządkowe. Zostanie teraz przeprowadzona analiza czynnikowa na podstawie macierzy

korelacji liniowej Pearsona, a następnie na podstawie macierzy korelacji polichorycznych.

Tabela 2. Liczba wypełnionych ankiet w ocenie osiemnastu kaw

Nr kawy	Nazwa kawy	Liczba ankiet	Nr kawy	Nazwa kawy	Liczba ankiet
1	Tchibo Mild	111	10	Mokate Cappucino	111
2	Tchibo Family	108	11	Celmar	93
3	Tchibo Mocca	75	12	Pedros Feinster	58
4	Tchibo Bistro	63	13	Mk-Cafe Natural	51
5	Tchibo Exclusive	79	14	Galaxia Sun	49
6	Jacobs Krönung	116	15	Maxwell House	105
7	Jacobs Merido	77	16	Nescafe Classic	96
8	Jacobs Maxima	50	17	Nescafe Gold	59
9	Jacobs Goldene Tasse	54	18	Mk-Cafe Premium	46

Źródło: zestawienie własne.

Tabela 3. Mediany ocen osiemnastu kaw dla dziewięciu badanych zmiennych przedstawionych na skali trzystopniowej

Nr kawy	Zmienna								
	<i>B</i>	<i>D</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>J</i>	<i>K</i>	<i>L</i>
1	2	2	2	2	3	3	2	2	2
2	2	2	2	2	3	3	2	2	3
3	2	2	2	2	3	3	3	2	2
4	2	2	2	2	3	3	2	2	2
5	2	2	2	2	3	3	2	2	2
6	1	3	2	2	3	3	3	2	2
7	2	2	2	2	3	3	2	2	2
8	2	2	2	2	3	3	2	2	2
9	2	2	2	3	3	3	2	2	2
10	2	2	3	3	3	3	2	3	2
11	3	2	2	2	2	3	2	2	2
12	2	2	2	2	2	3	2	2	2
13	2	2	2	2	2	3	2	2	2
14	2	2	2	2	2	2	2	2	2
15	2	2	3	3	3	3	2	3	2
16	2	2	3	3	3	3	2	3	2
17	2	2	3	3	3	3	3	3	2
18	2	2	2	2	2	3	2	2	2

Źródło: opracowanie własne na podstawie przeprowadzonego badania ankietowego.

W tabeli 4 zamieszczono macierz korelacji liniowych Pearsona, w tab. 5 – macierz korelacji polichorycznych, w tab. 6 zaś – wyniki analizy czynnikowej przeprowadzonej metodą głównych składowych z rotacją *quartimax* dla korelacji Pearsona oraz metodą osi głównych z rotacją *quartimax* dla korelacji polichorycznych.

Tabela 4. Macierz korelacji Pearsona dla zmiennych mierzonych na skali trzystopniowej

Zmienna	B	D	F	G	H	I	J	K	L
B	1,000	-0,728	0,000	0,000	-0,372	-0,000	-0,447	0,000	-0,000
C	-0,728	1,000	-0,130	-0,150	0,150	0,059	0,542	-0,130	-0,059
D	0,000	-0,130	1,000	0,862	0,331	0,130	0,120	1,000	-0,130
F	0,000	-0,150	0,862	1,000	0,385	0,150	0,055	0,862	-0,150
G	-0,372	0,150	0,331	0,385	1,000	0,391	0,277	0,331	0,150
H	-0,000	0,059	0,130	0,150	0,391	1,000	0,108	0,130	0,059
I	-0,447	0,542	0,120	0,055	0,277	0,108	1,000	0,120	-0,108
J	0,000	-0,130	1,000	0,862	0,331	0,130	0,120	1,000	-0,130
K	-0,000	-0,059	-0,130	-0,150	0,150	0,059	-0,108	-0,130	1,000
L	1,000	-0,728	0,000	0,000	-0,372	-0,000	-0,447	0,000	-0,000

Źródło: obliczenia własne z wykorzystaniem pakietu Statistica.

Tabela 5. Macierz korelacji polichorycznych

Zmienna	B	D	F	G	H	I	J	K	L
B	0,9965	-0,9965	0,0000	0,0000	-0,9911	0,0000	-0,9966	0,0000	0,0000
D	-0,9965	1,0000	-0,9565	-0,9680	0,9688	0,9344	0,9966	-0,9565	-0,9344
F	0,0000	-0,9565	1,0000	0,9998	0,9860	0,9684	0,2290	1,0000	-0,9565
G	0,0000	-0,9680	0,9998	1,0000	0,9878	0,9577	0,1072	0,9998	-0,9680
H	-0,9911	0,9688	0,9860	0,9878	1,0000	0,9934	0,9795	0,9860	0,9688
I	0,0000	0,9344	0,9684	0,9577	0,9934	1,0000	0,9502	0,9684	0,9344
J	-0,9966	0,9966	0,2290	0,1072	0,9795	0,9502	1,0000	0,2290	-0,9504
K	0,0000	-0,9565	1,0000	0,9998	0,9860	0,9684	0,2290	1,0000	-0,9565
L	0,0000	-0,9344	-0,9565	-0,9680	0,9688	0,9344	-0,9504	-0,9565	1,0000

Źródło: obliczenia własne z wykorzystaniem pakietu R.

Tabela 6. Ładunki czynnikowe uzyskane dla zmiennych mierzonych na skali trzystopniowej

Zmienna	Korelacje liniowe Pearsona			Korelacje polichoryczne		
	czynnik 1	czynniki 2	czynnik 3	czynnik 1	czynniki 2	czynniki 3
B	0,013	0,863	-0,084	0,022	0,724	0,047
D	-0,162	-0,889	-0,026	-0,458	-1,087	0,088
F	0,973	0,004	0,020	1,002	0,013	0,074
G	0,932	0,027	0,061	1,001	0,045	0,067
H	0,392	-0,364	0,654	0,600	-0,725	-0,631
I	0,158	-0,068	0,689	0,620	-0,540	-0,625
J	0,129	-0,763	0,002	0,248	-0,972	0,207
K	0,973	0,004	0,020	1,002	0,013	0,074
L	-0,250	0,156	0,664	-0,375	0,244	-1,163
% wyj. wariacji	30,45	22,80	13,56	41,69	35,31	22,08

Źródło: obliczenia własne z wykorzystaniem pakietu Statistica.

Wyniki analizy czynnikowej przeprowadzonej dla macierzy korelacji Pearsona różnią się od wyników dla korelacji polichorycznych, zatem dla zmiennych mie-

rzonych na skali trzystopniowej należy bazować na macierzy korelacji polichorycznych.

Na podstawie wyników z tab. 6 dotyczących korelacji polichorycznych otrzymano trzy czynniki opisane przez następujące zmienne:

- pierwszy czynnik jest scharakteryzowany przez następujące zmienne: wygodą przygotowania (*F*), czystość przygotowania (*G*) i funkcjonalność opakowania (*K*);
- drugi czynnik jest opisany przez zmienne: cena (*B*), aromat (*D*), estetyka opakowania (*H*) i rzetelność informacji na opakowaniach (*J*);
- trzeci czynnik: promocje (*L*).

Zmienna dostępność w sprzedaży (*I*) nie bierze istotnego udziału w opisie czynników.

Następnie te same zmienne przedstawiono na skali siedmiostopniowej (tab. 7) i ponownie przeprowadzono analizę czynnikową.

Tabela 7. Mediany ocen osiemnastu kaw dla dziesięciu badanych zmiennych przedstawionych na skali siedmiostopniowej

Nr kawy	Zmienna								
	<i>B</i>	<i>D</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>J</i>	<i>K</i>	<i>L</i>
1	4	4	5	5	6	6	5	5	4
2	5	4	5	5	6	7	5	5	5
3	4	5	5	5	6	6	6	5	4
4	4	4	5	5	6	6	5	5	4
5	3	5	5	5	6	6	5	5	4
6	2	6	5	5	6	7	6	5	5
7	5	4	5	5	6	6	5	5	5
8	4	4	5	5	6	6	5	5	4
9	3	4	5	6	6	6	5	5	4
10	5	4	6	6	6	7	5	6	4
11	7	3	4	4	4	6	4	4	3
12	5	4	5	5	5	6	5	5	4
13	5	4	5	5	5	6	5	5	4
14	4	4	5	4	4	5	4	4	3
15	4	5	6	6	6	6	5	6	4
16	4	4	6	6	6	6	5	6	5
17	3	4	6	6	6	6	6	6	4
18	5	4	5	5	5	6	5	4	4

Źródło: opracowanie własne na podstawie przeprowadzonego badania ankietowego.

W tabeli 8 zamieszczono macierz korelacji liniowych Pearsona między zmiennymi, w tab. 9 zaś wyniki analizy po zastosowaniu metody głównych składowych z rotacją *quartimax*.

Pogrubioną czcionką zaznaczono istotne ładunki czynnikowe większe co do wartości bezwzględnej od 0,65. Wyodrębnione czynniki wyjaśniają 76,07% zasobu zmienności wspólnej wszystkich zmiennych. Inne metody wyodrębniania czynni-

ków dały zbliżone wyniki, mianowicie zmienne mające istotny ładunek w prawie wszystkich metodach opisują ten sam czynnik.

Tabela 8. Macierz korelacji Pearsona

Zmienna	B	D	F	G	H	I	J	K	L
B	1,000	-0,725	-0,376	-0,398	-0,563	-0,050	-0,609	-0,349	-0,312
D	-0,725	1,000	0,236	0,196	0,463	0,300	0,637	0,253	0,399
F	-0,376	0,236	1,000	0,832	0,519	0,162	0,389	0,865	0,327
G	-0,398	0,196	0,832	1,000	0,704	0,336	0,500	0,868	0,435
H	-0,563	0,463	0,519	0,704	1,000	0,497	0,686	0,711	0,691
I	-0,050	0,300	0,162	0,336	0,497	1,000	0,437	0,369	0,595
J	-0,609	0,637	0,389	0,500	0,686	0,437	1,000	0,502	0,540
K	-0,349	0,253	0,865	0,868	0,711	0,369	0,502	1,000	0,456
L	-0,312	0,399	0,327	0,435	0,691	0,595	0,540	0,456	1,000

Źródło: obliczenia własne z wykorzystaniem pakietu SPSS.

Tabela 9. Ładunki czynnikowe uzyskane za pomocą metody głównych składowych po rotacji *quartimax*

Zmienna	Korelacje liniowe Pearsona			Korelacje polichoryczne		
	czynnik 1	czynnik 2	czynnik 3	czynnik 1	czynnik 2	czynnik 3
B	-0,300	-0,891	0,086	-0,313	-0,929	0,115
D	0,063	0,892	0,203	0,272	0,901	0,288
F	0,934	0,132	-0,038	0,972	0,178	-0,026
G	0,927	0,131	0,177	0,965	0,117	0,157
H	0,650	0,449	0,491	0,721	0,486	0,488
I	0,165	0,057	0,899	0,334	0,047	0,915
J	0,359	0,680	0,402	0,505	0,654	0,460
K	0,927	0,126	0,216	0,953	0,110	0,224
L	0,312	0,292	0,754	0,452	0,278	0,778
% wyj. wariacji	33,06	23,93	19,08	40,48	24,77	20,63

Źródło: obliczenia własne z wykorzystaniem pakietu Statistica.

W tabeli 10 zamieszczono macierz korelacji polichorycznych między tymi zmiennymi.

Tabela 10. Macierz korelacji polichorycznych

Zmienna	B	D	F	G	H	I	J	K	L
B	0,957	-0,879	-0,435	-0,425	-0,735	-0,024	-0,685	-0,380	-0,306
D	-0,879	0,957	0,457	0,361	0,963	0,395	0,855	0,429	0,584
F	-0,435	0,457	0,998	0,998	0,985	0,262	0,586	0,997	0,495
G	-0,425	0,361	0,998	0,999	0,988	0,513	0,678	0,999	0,595
H	-0,735	0,963	0,985	0,988	1,000	0,981	0,981	0,985	0,985
I	-0,024	0,395	0,262	0,513	0,981	0,998	0,662	0,562	0,829
J	-0,685	0,855	0,586	0,678	0,981	0,662	0,998	0,687	0,736
K	-0,380	0,429	0,997	0,999	0,985	0,562	0,687	0,999	0,619
L	-0,306	0,584	0,495	0,595	0,985	0,829	0,736	0,619	0,999

Źródło: obliczenia własne z wykorzystaniem pakietu R.

W praktyce macierz korelacji polichorycznych często nie jest określona dodatnio, tzn. że posiada jedną lub więcej ujemnych wartości własnych, a co za tym idzie – nie ma macierzy odwrotnej. Nie jest to problem sam w sobie, jednak należy pamiętać, że niektóre metody wyodrębniania czynników wymagają macierzy odwrotnej. Do tych metod zaliczamy metodę największej wiarygodności i uogólnioną metodę najmniejszych kwadratów. Natomiast metody, które nie wymagają macierzy odwrotnej, to metody nieważonych najmniejszych kwadratów, osi głównych oraz głównych składowych.

Dostępne są programy, np. TeMat, które „wygładzają” macierz nieokreśloną dodatnio. Polega to na usuwaniu z macierzy składników, które powodują ujemne wartości własne. Tak poprawiona macierz korelacji prowadzi do bardziej dokładnych estymacji czynników, jednakże nadal nie jest odpowiednia do metody największej wiarygodności i uogólnionej metody najmniejszych kwadratów.

Macierz korelacji w tab. 10 ma trzy ujemne wartości własne, „wygładzoną” macierz zamieszczono w tab. 11.

Tabela 11. „Wygładzona” macierz korelacji polichorycznych

Zmienna	B	D	F	G	H	I	J	K	L
B	1,000	-0,878	-0,460	-0,406	-0,624	-0,055	-0,691	-0,375	-0,314
D	-0,878	1,000	0,438	0,386	0,779	0,405	0,828	0,436	0,591
F	-0,460	0,438	1,000	0,937	0,779	0,295	0,572	0,938	0,492
G	-0,406	0,386	0,937	1,000	0,830	0,472	0,658	0,950	0,587
H	-0,624	0,779	0,779	0,830	1,000	0,714	0,890	0,846	0,842
I	-0,055	0,405	0,295	0,472	0,714	1,000	0,613	0,548	0,810
J	-0,691	0,828	0,572	0,658	0,890	0,613	1,000	0,651	0,728
K	-0,375	0,436	0,938	0,950	0,846	0,548	0,651	1,000	0,608
L	-0,314	0,591	0,492	0,587	0,842	0,810	0,728	0,608	1,000

Źródło: obliczenia własne z wykorzystaniem pakietu TeMat.

W tabeli 9 przedstawiono wyniki analizy po zastosowaniu metody głównych składowych z rotacją *quartimax*.

Wyodrębnione czynniki wyjaśniają 85,88% zasobu zmienności wspólnej wszystkich zmiennych.

Wyniki analizy czynnikowej przeprowadzonej dla korelacji liniowych Pearsona i polichorycznych są zbliżone oraz prowadzą do podobnych wniosków:

- pierwszy czynnik jest scharakteryzowany przez takie zmienne, jak: wygoda przygotowania (*F*), czystość przygotowania (*G*), estetyka opakowania (*H*) i funkcjonalność opakowania (*K*); czynnik ten nazwano atrakcyjnością opakowania;
- drugi czynnik jest opisany przez zmienne: cena (*B*), aromat (*D*) i rzetelność informacji na opakowaniach (*J*); nazwano go właściwościami smakowymi;
- trzeci czynnik opisują następujące zmienne: dostępność w sprzedaży (*I*) i promocje (*L*); nazwano go powszechnością kawy.

Czynniki te nie pokrywają się z czynnikami otrzymanymi dla zmiennych mierzonych na skali trzystopniowej, ponieważ tam nastąpiła znaczna utrata informacji.

W literaturze zaleca się stosowanie korelacji liniowych Pearsona dla zmiennych mierzonych na skalach, które mają sześć i więcej stopni. Zaprezentowany przykład potwierdził, że w odniesieniu do zmiennych mierzonych na skali siedmiostopniowej analiza czynnikowa przeprowadzona dla korelacji liniowych Pearsona i polichorycznych daje podobne wyniki.

3. Podsumowanie

Przeprowadzając analizę czynnikową dla zmiennych porządkowych, należy pamiętać o następujących wskazówkach:

1. Korelacje liniowe Pearsona można stosować w odniesieniu do zmiennych mierzonych na skalach, które posiadają sześć i więcej kategorii, ponieważ prowadzą do podobnych wyników jak korelacje polichoryczne.

2. W przypadku zmiennych mierzonych na skalach, które mają pięć i mniej kategorii, należy bazować na macierzy korelacji polichorycznych (dla zmiennych dychotomicznych wyznacza się korelacje tetrachoryczne).

3. Jeśli macierz korelacji polichorycznych nie jest określona dodatnio, to jako metodę wyodrębniania czynników należy wybrać np. metodę nieważonych najmniejszych kwadratów, osi głównych lub głównych składowych. Wskazane jest w tym wypadku „wygładzenie” macierzy korelacji.

Literatura

- Bartholomew D.J. (1980), *Factor Analysis for Categorical Data*, „Journal of the Royal Statistical Society” Series B (Methodological) vol. 42, nr 3, s. 293-321.
- Bartholomew D.J., Steele F., Moustaki I., Galbraith J.I. (2002), *The Analysis and Interpretation of Multivariate Data for Social Scientists*, Chapman & Hall/CRC.
- Duffy D., Battistutta D., Mathews J., Martin N. (1998), *A Factor Analysis of Associations among Self-Reported Immune Related Symptoms in a Large Twin Sample*, „Twin Research” 1, s. 71-77.
- Kubinger K.D. (2003), *On Artificial Results Due to Using Factor Analysis for Dichotomous Variables*, „Psychology Science” vol. 45, s. 106-110.
- Kukuk M. (1998), *Analyzing Ordered Categorical Data derived from Elliptically Symmetric Distributions*, <http://w210.ub.uni-tuebingen.de/dbt/volltexte/2005/2135/pdf/156.pdf>.
- Konarski R. (2004), *Model cechy latentnej w analizie psychometrycznej testów I pozycji testowych*, [w:] B. Niemiecko, H. Szaleniec (red.), *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*, Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Ming Tan, Yinsheng Qu, J. Sunil Rao (1999), *Robustness of the Latent Variable Model for Correlated Binary Data*, „Biometrics” vol. 55, nr 1, s. 258-263.
Olivares A. M., (2003), *Testing Categorized Bivariate Normality with Two-Stage Polychoric Correlation Estimates*, <http://www.ub.es/gdne/documents/bivariate%20normality%20wp.pdf>.

FACTOR ANALYSIS FOR ORDERED CATEGORICAL VARIABLES

Summary

The method of factor analysis is widely used as an exploratory tool to reduce the dimensionality of multivariate data. The fact that the standard model is strictly applicable only when the manifest variables are scaled is a serious limitation in economic science where the variables are often categorical.

The classical factor analysis is based at Pearson correlation, while factor analysis for ordered variables uses polychoric correlation or tetrachoric for binary data.

The article presents the usage of factor analysis to reduce the set of coffee characteristics in order to reach the few major factors influencing consumers' choice of a brand.